

一种基于 DCT 的频域运动估值算法

曹宁 吴敏

(河海大学计算机及信息工程学院, 常州 213022)

摘要 用软件实现在低码率信道上的实时视频编解码是图象压缩领域的研究热点, 本文就此提出了一种基于离散余弦变换(DCT)系数的频域块匹配法, 通过引入阈值矩阵, 降低了运动估值的运算量, 提高了搜索效率. 计算机仿真表明, 该算法可在 PC 机上用软件实时实现, 具有较好的实用价值.

关键词 H. 263 频域块匹配法 阈值矩阵

中图分类号: TN919.81 **文献标识码:** A **文章编号:** 1006-8961(2002)02-0125-03

A DCT Based Motion Estimation in Frequency Domain

CAO Ning, WU Min

(College of Computer & Information Engineering Hohai University, Changzhou 213022)

Abstract To realize the real-time video coding in low bit rate communication by PC software is the hotspot in image compression field, thus a frequency domain block matching algorithm based on discrete cosine transform (DCT) coefficients is proposed in this paper. By utilizing threshold matrix, the motion estimation computation is greatly reduced and the time used in searching motion vectors is saved. The DCT-based frequency domain block matching algorithm can encode and decode motion pictures quickly and efficiently. Its reconstructed picture quality is higher than the 3SS algorithm and is closer to the FS algorithm. Simulation results show that the proposed DCT based block matching algorithm and the video coding scheme has high performance and is easy for implementation. Because of its low computing complexity, it can be real-time realized by PC software and has a good value in practice.

Keywords H. 263, Frequency domain block matching algorithm, Threshold matrix

0 引言

ITU-T 于 1996 年 3 月正式通过了码率低于 64Kb/s 的低比特率信道传输的 H. 263 建议^[1]. 该建议是 ITU-T H. 324 低比特率多媒体通信终端建议中视频编码的核心部分, 其应用目标是基于 PSTN 或 ISDN 网的可视电话(Videophone)和会议电视(Conference Television)及无线通信中的视频传输, 它的出现使得在带宽有限的信道上传送质量较好的视频图象成为可能.

H. 263 的视频信源编码方法与 ITU-T H. 261

建议相类似, 采用 DCT/MC 的混合编码方法, 其视频编码器的复杂性主要取决于运动估值算法的复杂度. 目前, 一些典型的快速算法有三步搜索法(3SS)^[2]、四步搜索法(4SS)^[3]、二维对数搜索法(LOGS)^[4]和动态搜索窗调整搜索法(DSWDS)^[5]等, 这些运动估值算法都是基于块匹配准则的, 是在空间域上直接进行的运动估值. 由 DCT/MC 编码框图(图 1)可知, 在此方案中, 编码器要比解码器多做一次 DCT 逆变换和反量化运算, 因而编解码的运算量是不平衡的. 为了实现在低码率信道上的视频传输以及用软件的实时、高效视频编码, 提出了一种基于 DCT 的块匹配运动估值法.

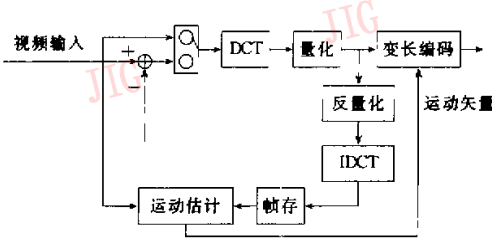


图1 DCT/MC 视频编码框图

1 基于 DCT 的频域块匹配法

1.1 频域块匹配法原理

DCT 变换作为一种准最佳变换,其系数矩阵具有以下性质:

(1) 矩阵中各变换系数间几乎不相关.

(2) 系数矩阵的能量主要集中在矩阵的左上角,即直流分量和低频区.

因此,可以利用 DCT 变换的这些性质,在频域上使用块匹配法进行运动估值.其基本思想是:经 DCT 变换后,系数幅值较小的部分为高频部分,即图象运动剧烈的地方.对可视电话而言,图象的运动相对较为缓慢,因而其低频成分相对较多,高频成分相对较少.利用这一特性,可在搜索域中引入一阈值矩阵,该矩阵的作用是对 DCT 系数矩阵进行过滤,使得搜索仅在运动相对剧烈的高频部分进行,从而减少了搜索量,提高了块匹配速度.

从图 2 的编码框图也可看出,基于频域的块匹配运动估值编码器只需做一次 DCT 变换和量化,省下了做 DCT 逆变换和反量化的步骤,不仅可提高编码效率,且使得编解码器的运算量基本平衡.

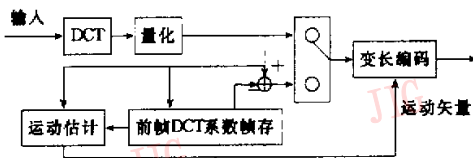
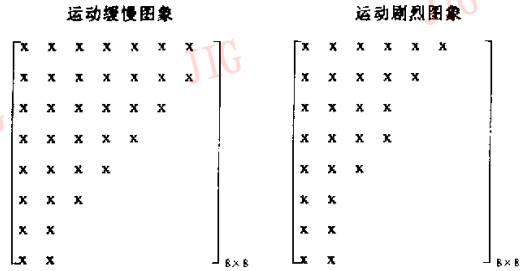


图2 频域运动估值编码框图

1.2 阈值矩阵的选取

本方案的关键是对阈值矩阵的选取.所选取的矩阵应满足以下条件:舍弃系数值大于某一给定值 ϵ 的 DCT 系数,而仅保留那些小于 ϵ 的系数.对不同图象而言, ϵ 的取值都不尽相同,而本文是用一阈值矩阵对 DCT 系数矩阵进行加权的方法来实现的.

通过对不同运动程度图象的 DCT 系数矩阵的研究发现,对于运动较缓慢的图象和运动较剧烈的图象,其大系数数据集中分布的情况分别如下(x 表示大系数所在的位置):



也就是说,大系数的分布在水平和垂直方向并不是对称的,因而加权矩阵不宜采用简单的水平、垂直或对角线对称的矩阵.为更好地适应不同运动程度图象的情况,对于搜索域为 $(\pm 16, \pm 16)$ 采用如下矩阵:

$$P = \begin{bmatrix} A & \dots & A \\ A & \dots & A \\ \vdots & \vdots & \vdots \\ A & \dots & A \end{bmatrix}_{8 \times 8}$$

其中,

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

将该矩阵对搜索区域中的各系数进行加权处理,即可滤除低频分量.

1.3 具体的搜索策略

通常,视频序列图象中宏块的运动矢量在搜索区域内不是均匀分布的.图象序列“Miss American”和“Table Tennis”分别有 83% 和 81% 的运动位移分布在以估值点为中心的 3×3 区域内,由此可见,运动矢量在搜索域中有很强的空间分布特性——中心偏置分布特性.针对这一特性,本文采用搜索步长为 1 的一种搜索法,其具体步骤如下:

(1) 以估值初始点为中心,取搜索步长为 1,计算其 3×3 搜索窗上 9 个点的匹配函数值.若当前搜索窗的中心点为匹配函数值最小的点,则中止搜索

转第(4)步;否则,以当前匹配函数值最小的点作为下一次搜索窗的中心点。

(2)以新的搜索窗中心点为中心,取搜索步长为1,计算其 3×3 搜索窗上9个点的匹配函数值。若当前搜索窗的中心点为匹配函数值最小的点,则中止搜索转第(4)步;否则,以当前匹配函数值最小的点作为下一次搜索窗的中心点。

(3)重复第(2)步,直至搜索窗的中心点到达搜索区域的边界为止。

(4)以找到的最佳整数像素匹配点为中心,取搜索步长为 $1/2$,进行一次半像素全搜索,得到最终的运动矢量。

2 计算机模拟实验结果

采用QCIF格式的标准图象序列Miss American(图象较简单,运动较缓慢)和Carphone(图象相对复杂,运动较剧烈)各50帧,在P166计算机上分别对全搜索法(FS)、三步搜索法(3SS)及本文提出的频域块匹配法(FBMA)进行了对比实验。实验采用MAD(帧间绝对差)准则^[6]作为块匹配准则,实验结果见表1、表2。

表1 搜索区域为 $(\pm 5, \pm 5)$ 时的实验结果

运动估值 算法	平均搜索次数		平均编码时间(s)		PSNR(dB)	
	Miss 序列	Carphone 序列	Miss 序列	Carphone 序列	Miss 序列	Carphone 序列
FS	95	95	0.396	0.408	36.41	34.29
3SS	15.41	15.22	0.127	0.132	35.76	33.58
FBMA	17.34	15.93	0.128	0.133	36.12	34.06

表2 搜索区域为 $(\pm 16, \pm 16)$ 时的实验结果

运动估值 算法	平均搜索次数		平均编码时间(s)		PSNR(dB)	
	Miss 序列	Carphone 序列	Miss 序列	Carphone 序列	Miss 序列	Carphone 序列
FS	975	975	0.531	0.541	36.59	34.37
3SS	44.26	42.87	0.193	0.198	35.87	33.71
FBMA	21.37	18.65	0.133	0.136	36.25	34.15

从以上实验数据可看出,本文所提出的频域块匹配法性能较好,尤其在搜索域较大时,其平均搜索次数,编码时间要明显优于常用的3SS法,重建图象质量也很接近于FS算法。由此也证明了该算法确实能减少搜索次数,提高编码效率。

3 程序的优化

采用频域块匹配法后,编码速度有了明显提高,为了更好地在PC机上用软件进行实时编解码,使用汇编语言对耗时较大的DCT运算进行了程序代码的优化。通过直接对CPU进行控制,充分利用了通用寄存器,尽量减少浮点和乘除运算的次数及对内存变量的频繁调用。经过优化,在搜索域为 $(\pm 16, \pm 16)$ 时,Miss American序列和Carphone序列的平均编码时间分别降低为0.109s和0.112s,基本满足实时编解码的要求。

4 结论

本文提出的基于DCT的频域块匹配法,不仅减少了运动估值的运算量,且还使得编解码器的运算量基本平衡,较大地提高了编码效率。经过对程序的优化,实现了在PC机上用软件进行的实时编解码,具有一定的实用价值。

参考文献

- 1 ITU-T Recommendation H.263-1996.3, Video coding for low bit rate communication[S].
- 2 Koya T, Iimura K, Hirano A *et al.* Motion compensated interframe coding for video conferencing[A]. In: Proc. NTC81[C]. New Orleans, LA, 1981, C9. 6. 1~9. 6. 5.
- 3 Po M, Ma W C. A novel four-step algorithm for fast block motion estimation[J]. IEEE Trans. CASVT, 1996, 6(3): 313~317.
- 4 Jain J R, Jain A K. Displacement measurement and its application in interframe image coding[J]. IEEE Trans, Commun, 1981, COM-29(12): 1799~1808.
- 5 Lee W, Wang J F, Lee J Y *et al.* Dynamic search-window adjustment and interlaced search for block-matching algorithm[J]. IEEE Trans, CASVT, Feb. 1993, 3(1): 85~87.
- 6 吴乐南. 数据压缩的原理与应用[M]. 北京: 电子工业出版社, 1995.

曹宁 1962年生,副教授,从事多媒体信息处理及传输技术的研究。

吴敏 1976年生,硕士,从事视频编码技术的研究。